_S9C_
_10/09_ **OIPE**

ENTERED

```
            RAW SEQUENCE LISTING              DATE:  6-15-2001
            PATENT APPLICATION:  US/09/851,410   TIME: 16:06:06

            Input Set : N:\Crf3\RULE60\09851410.txt
            Output Set: N:\CRF3\06152001\I851410.raw


            SEQUENCE LISTING
     3  (1) GENERAL INFORMATION:
     5      (i) APPLICANT: Reyes, Gregory R
     6                     Yarbough, Patrice O
     7                     Bradley, Daniel W
     8                     Krawczynski, Krzysztof Z
     9                     Tam, Albert
    10                     Fry, Kirk E
    11     (ii) TITLE OF INVENTION: DNA Sequences of Enterically Transmitted
    13                      Non-A/Non-B Hepatitis Viral Agent
    15    (iii) NUMBER OF SEQUENCES: 20
    17     (iv) CORRESPONDENCE ADDRESS:
    18          (A) ADDRESSEE: Dehlinger & Associates
    19          (B) STREET: 350 Cambridge Avenue, Suite 250
    20          (C) CITY: Palo Alto
    21          (D) STATE: CA
    22          (E) COUNTRY: USA
    23          (F) ZIP: 94306
    24     (v) COMPUTER READABLE FORM:
    26          (A) MEDIUM TYPE: Floppy disk
    27          (B) COMPUTER: IBM PC compatible
    28          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    29          (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
    31     (vi) CURRENT APPLICATION DATA:
C--> 32          (A) APPLICATION NUMBER: US/09/851,410
C--> 33          (B) FILING DATE: 07-May-2001
    35    (vii) PRIOR APPLICATION DATA:
    36          (A) APPLICATION NUMBER: 09/128,275
    37          (B) FILING DATE:
    40          (A) APPLICATION NUMBER: US 07/681,078
    41          (B) FILING DATE: 05-APR-1991
    44          (A) APPLICATION NUMBER: US 07/505,888
    45          (B) FILING DATE: 05-APR-1990
    48          (A) APPLICATION NUMBER: US 07/421,921
    49          (B) FILING DATE: 13-OCT-1989
    52          (A) APPLICATION NUMBER: US 07/367,486
    53          (B) FILING DATE: 16-JUN-1989
    56          (A) APPLICATION NUMBER: US 07/480,070
    57          (B) FILING DATE: 11-APR-1989
    60          (A) APPLICATION NUMBER: US 07/161,487
    61          (B) FILING DATE: 17-JUN-1988
    63   (viii) ATTORNEY AGENT INFORMATION:
    64          (A) NAME: Britthery, Joann B.
    65          (B) REGISTRATION NUMBER: 41,49
    66          (C) REFERENCE DOCKET NUMBER: 4140-1444.43
    68     (ix) TELECOMMUNICATION INFORMATION:
    69          (A) TELEPHONE: 650-324-4445
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/851,410

DATE:
TIME:

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

```
70          (B) TELEFAX:  (651) 324-1301
72 (2) INFORMATION FOR SEQ ID NO: 1:
74      (i) SEQUENCE CHARACTERISTICS:
75          (A) LENGTH: 1295 base pairs
76          (B) TYPE: nucleic acid
77          (C) STRANDEDNESS: double
78          (D) TOPOLOGY: linear
W--> 80     (ii) MOLECULE TYPE: DNA
82     (iii) HYPOTHETICAL: NO
84      (iv) ANTI-SENSE: NO
86      (vi) ORIGINAL SOURCE:
87          (C) INDIVIDUAL ISOLATE: 1.35 kb EcoRI insert of ET1.1,
88                                   forward sequence
90      (ix) FEATURE:
91          (A) NAME/KEY: CDS
92          (B) LOCATION: 1..1293
94      (ix) FEATURE:
95          (A) NAME/KEY: CDS
96          (B) LOCATION: 2..1294
97      (ix) FEATURE:
98          (A) NAME/KEY: CDS
100         (B) LOCATION: 3..1295
103     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
105 AGACCTGTCC CTGTTGCAGC TGTTCTACCA CCCTGCCCCG AGCTCGAACA GGGCCTTCTC      60
107 TACCTGCCCC AGGAGCTCAC CACCTGTGAT AGTGTCGTAA CATTTGAATT AACAGACATT     120
109 GTGCACTGCC GCATGGCCGC CCCGAGCCAG CGCAAGGCCG TGCTGTCCAC ACTCGTGGGC     180
111 CGCTACGGCG GTCGGCACAAA GCTCTACAAT GCTTCCCACT CTGATGTTCG CGACTCTCTC     240
113 GCCCGTTTTA TCCCGGCCAT TGGCCCCGTA CAGGTTACAA CTTGTGAATT GTACGAGCTA     300
115 GTGGAGGCCA TGGTCGAGAA GGGCCAGGAT GGCTCCGCCG TCCTTGAGCT TGATCTTTGC     360
117 AACGGTGACG TGTCCAGGAT CACCTTCTTC CAGAAAGATT GTAACAAGTT CACCACAGGT     420
119 GAGACCATTG CCCATGGTAA AGTGGGCCAG GGCATCTCGG CCTGGAGCAA GACCTTCTGC     480
121 GCCCTCTTTG GCCCTTGGTT CCGGCGTATT GAGAAGGCTA TTCTGGCGCT GCTCCCTCAG     540
123 GGTGTGTTTT ACGGTGATGC CTTTGATGAC ACCGTCTTCT CGGCGGCTGT GGCGGCAGCA     600
125 AAGGCATCCA TGGTGTTTGA GAATGACTTT TCTGAGTTTG ACTCCACCCA GAATAACTTT     660
127 TCTCTGGGTC TAGAGTGTGC TATTATGGAG GAGTGTGGGA TGCCGCAGTG GCTCATCCGC     720
129 CTGTATCACC TTATAAGGGC TGCGTGGATC TTGCAGGCGC CGAAGGAGTC TCTGCGAGGG     780
131 TTTTGGAAGA AACACTCCGG TGAGCCTGGC ACTCTTCTAT GGAATACTGT CTGGAATATG     840
133 GCCGTTATTA CCCACTGTTA TGACTTCCGC GATTTTCAGG TGGCTGCCTT TAAAGGTGAT     900
135 GATTCGATAG TGCTTTGCAG TGAGTATGGT CAGAGTCCAG GAGCTGCTGT CCTGATCGCC     960
137 GGCTGTGGCT TGAAGTTGAA GGTAGAGTTTC CGCGCGATCG GTTTGTATGC AGGTGTTGTG     1020
139 GTGGCCCCCG GCCTTGGCGC GCTTCCTGAT GTTGTGGCGCT TGGCGGCCG GCTTACCGAG     1080
141 AAGAATTGGG GCCCTGGCCC TGAGCGGGCG GAGCAGCTCG GCCTCGCTGT TAGTGATTTC     1140
143 CTCCGCAAGC TCACGAATGT AGCTCAGATG TGTGTGATG TTCTTCCCG TGTTTATGCG     1200
145 GTTTCCCCTG GACTCGTTCA TAACCTGATT GGCATCGTAT AGCGTGTTGC TTATGGCAAG     1260
147 GCACATTTCA CTGAGTCAGT AAAACCAGTG CTCCA                              1295
150 (2) INFORMATION FOR SEQ ID NO: 2:
152     (i) SEQUENCE CHARACTERISTICS:
153         (A) LENGTH: 431 amino acids
154         (B) TYPE: amino acid
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/851,410

DATE: ...

TIME: ...

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

```
155              TOPOLOGY: linear
157      (ii) MOLECULE TYPE: protein
159      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
161  Arg Pro Val Pro Val Ala Ala Val Leu Ser Pro Tyr Pro Glu Leu Glu
162    1                5                                      15
164  Gln Gly Leu Leu Tyr Leu Pro Gln Glu Leu Thr Thr Tyr Asp Ser Val
16          20                      25                      30
166  Val Thr Phe Glu Leu Thr Asp Ile Val His Lys Arg Met Ala Ala Pro
168          35                      40                      45
170  Ser Gln Arg Lys Ala Val Leu Ser Thr Leu Val Gly Arg Tyr Gly Gly
171      50                      55                      60
173  Arg Thr Lys Leu Tyr Asn Ala Ser His Ser Asp Val Arg Asp Ser Leu
174  65                      70                      75                   80
176  Ala Arg Phe Ile Pro Ala Ile Gly Pro Val Gln Val Thr Thr Cys Glu
177                  85                      90                      95
179  Leu Tyr Glu Leu Val Glu Ala Met Val Glu Lys Gly Gln Asp Gly Ser
180                  100                     105                     110
181  Ala Val Leu Glu Leu Asp Leu Cys Asn Arg Asp Val Ser Arg Ile Thr
182          115                     120                     125
184  Phe Phe Gln Lys Asp Cys Asn Lys Phe Thr Thr Gly Glu Thr Ile Ala
186          130                     135                     140
188  His Gly Lys Val Gly Gln Gly Ile Ser Ala Trp Ser Lys Thr Phe Cys
189  145                     150                     155                    160
191  Ala Leu Phe Gly Pro Trp Phe Arg Ala Ile Glu Lys Ala Ile Leu Ala
192                  165                     170                     175
194  Leu Leu Pro Gln Gly Val Phe Tyr Gly Asp Ala Phe Asp Asp Thr Val
195                  180                     185                     190
197  Phe Ser Ala Ala Val Ala Ala Ala Lys Ala Ser Met Val Phe Glu Asn
198                  195                     200                     205
200  Asp Phe Ser Glu Phe Asp Ser Thr Gln Asn Asn Phe Ser Leu Gly Leu
201       210                     215                     220
202  Glu Cys Ala Ile Met Glu Glu Cys Gly Met Pro Gln Trp Leu Ile Arg
204  225                     230                     235                    240
206  Leu Tyr His Leu Ile Arg Ser Ala Trp Ile Leu Gln Ala Pro Lys Glu
207                  245                     250                     255
209  Ser Leu Arg Gly Phe Trp Lys Lys His Ser Gly Glu Pro Gly Thr Leu
210                  260                     265                     270
211  Leu Trp Asn Thr Val Trp Asn Met Ala Val Ile Thr His Cys Tyr Asp
212                  275                     280                     285
213  Phe Arg Asp Phe Gln Val Ala Ala Phe Lys Gly Asp Asp Ser Ile Val
214       290                     295                     300
216  Leu Cys Ser Glu Tyr Arg Gln Ser Pro Gly Ala Ala Val Leu Ile Ala
217  305                     310                     315                    320
221  Gly Cys Gly Leu Lys Leu Lys Val Asp Phe Arg Pro Ile Gly Leu Tyr
22                  325                     330                     335
224  Ala Gly Val Val Val Ala Pro Gly Leu Gly Ala Leu Pro Asp Val Val
22                  340                     345                     350
227  Arg Phe Ala Gly Arg Leu Thr His Lys Asn Trp Gly Pro Gly Pro Gly
228       355                     360                     365
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/851,410

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

230 Arg Ala Glu Gly Leu Arg Leu Ala Val Thr Asp Phe Leu Arg Lys Leu
231   370   375

263 Thr Asn Val Ala Gln Met Cys Val Asp Val Val Ser Arg Val Tyr Gly
    385   390

Val Ser Pro Gly Leu Val His Asn Leu Ile Gly Met Leu Gln Ala Val
              405   410   415

Ala Asp Gly Lys Ala His Phe Thr Glu Arg Val Lys Pro Val Leu
    420   425   430

(2) INFORMATION FOR SEQ ID NO: 3:
     (i) SEQUENCE CHARACTERISTICS:
          A) LENGTH: 18 base pairs
          B) TYPE: nucleic acid
          C) STRANDEDNESS: single
          D) TOPOLOGY: linear
W--> 251   (ii) MOLECULE TYPE: DNA
     (iii) HYPOTHETICAL: NO
     (iv) ANTI-SENSE: NO
     (vi) ORIGINAL SOURCE:
          C) INDIVIDUAL ISOLATE: linker - top (5') sequence
     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
GGAATTCGCG CCCGCTCG   18
(2) INFORMATION FOR SEQ ID NO: 4:
     (i) SEQUENCE CHARACTERISTICS:
          A) LENGTH: 20 base pairs
          B) TYPE: nucleic acid
          C) STRANDEDNESS: single
          D) TOPOLOGY: linear
W--> 273   (ii) MOLECULE TYPE: DNA
     (iii) HYPOTHETICAL: NO
     (iv) ANTI-SENSE: NO
     (vi) ORIGINAL SOURCE:
          C) INDIVIDUAL ISOLATE: linker - bottom (3') sequence
     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
CGAGCGGGCG CGAATTCCTT   20
(2) INFORMATION FOR SEQ ID NO: 5:
     (i) SEQUENCE CHARACTERISTICS:
          A) LENGTH: 1235 base pairs
          B) TYPE: nucleic acid
          C) STRANDEDNESS: double
          D) TOPOLOGY: linear
W--> 294   (ii) MOLECULE TYPE: DNA
     (iii) HYPOTHETICAL: NO
     (iv) ANTI-SENSE: NO
     (vi) ORIGINAL SOURCE:
          C) INDIVIDUAL ISOLATE: 1.33 kb EcoRI insert of ET1.1,
                              reverse sequence
     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
TCGAGCACTG GTTTTATTGA CTCAGTGAAA TGT   TTTT TATCATAAAT AAACTCTAAT   60
ATGCCAATCA TTTATGAAT CACTCTAGAA AAA   TTAT AAACATGGA AAAACATT   120

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/851,410

DATE: [illegible]
TIME: [illegible]

Input Set : N:\Crf3\RULE60\09851410.txt
Output Set: N:\CRF3\06152001\I851410.raw

```
311 ACACACATCT CAGGTACATT CGTCAGCTTG CGGATGAAAT CACTAACAGC GAGGCGGAGC   180
312 TGTTTTGGTCG CGTCAGGGCC AGGGCCCCAA TTCTTCTTGG TAAGCCGGCC CCCGAAGCGC  240
31  ACAACATTAG GCAGCGCGCC AAGGCCGGGG GCCACCACAA CAGCTGCATA CAAACCGATC   300
31  GGGCCGAAAT CTACGTTCAA CTTCAAGCCA CAGCCGGCGA TCAGGACAGC AGTTCCTGGA   360
314 CTCCGAGCAT ACTCACTGCA AAGCACTATC GAATCATCAC CTTTAAAGGC AGCCACCTGA   420
321 AAATTGGCGGA ACTCATAACA GTGGGTAATA ACGGCCATAT TCCAGACAGT ATTCCATAGA   480
323 AGAGTGGCGG GCTCACCGGA GTGTTCTTC CAAAGCCTC GCAGAGACTC CTTCGGGGCC    540
32  TGCAAGATCC ACGCAGACCT TATAAGGTGA TACAGGCGGA TGAGCCACTG CGGCATCCCA   600
317 CATTTCTTCA CAATAGCACA CTCTCACCCC AGAGAAAAGT TATTCTCGGT GGAGTCAAAC   660
319 TCAGAAAAGT TATTCTCAAA CACCATGGAT GCCTTTGCTG CGGCCACAGC CGCCGAGAAG   720
3 1 ACGGTGTTAT TAAAGGCATC ACGGTAAAAC ACACGGTCAG GGAGCAGGGC CAGAATAGCC  780
323 TTCTCAATAG TGCGGAACCA AGGGCTAAAG AGGGCGCAGA AGGTCTTGCT CCAGGCCGAG   840
3 6 ATGCCGTCGG CCACTTTACC ATGGGTAATG GTCTCACCTG TGGTGAACTT GTTACAATCT   900
3 7 TTCTGCAAGA AGGTGATGCT GGACATGTCA CGGTTGCAAA GATCAAGCTC AAGGACGGCG   960
339 GAGCCATCCT GGCCCTTCTC GACCATGGCC TCCACTAGCT CGTACAATTC ACAAGTTGTA  1020
341 ACCTGTACGG GGCCAATGGC CGGGATAAAA CGGGCGAGAG AGTCGCGAAC ATCAGAGTGG  1080
343 GAAGCATTGT AGAGCTTTGT GCGACCGCCG TAGCGGCCCA CGAGTGTGGA CAGCACGGCC  1140
345 TTGCGGCTGGC TCGGGGCGGC CATGCGGCAG TGCACAATGT CTGTTAATTC AAATGTTACG  1200
347 ACACTATCAC AGGTGGTGAG TTCCTGGGGC AGGTAGAGAA GGCCCTGTTC GAGCTCGGGG  1260
349 CAGGGTCGTA GAACAGCTGG AACAGGCACA GGTCT                              1295
```

3 2 (2) INFORMATION FOR SEQ ID NO: 6:
3 4   (i) SEQUENCE CHARACTERISTICS:
3 5   (A) LENGTH: 7195 base pairs
3 6   (B) TYPE: nucleic acid
W--> 367   **(C) STRANDEDNESS: HEV - Burma strain**
3 8   (D) TOPOLOGY: linear
W--> 360 **(ii) MOLECULE TYPE: DNA**
3 2   (iii) HYPOTHETICAL: NO
3 4   (iv) ANTI-SENSE: NO
3 6   (vi) ORIGINAL SOURCE:
3 9   (ix) FEATURE:
3 0   (A) NAME/KEY: CDS
371   (B) LOCATION: 28..5106
3 3   (ix) FEATURE:
374   (A) NAME/KEY: CDS
3 9   (B) LOCATION: 5147..7126
3 7   (ix) FEATURE:
3 9   (A) NAME/KEY: CDS
3 4   (B) LOCATION: 5106..5474
3 2   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
 4 4 ATGGAGACCA CATATGTGGT TGATGCCATG GAGGCCCATC AGTTTATTAA GGCTCCTGGC   60
 6 ATCACTACTG CTATTGAGCA GGCTGTTCTA GCAGCGGCCA ACTCTGCCCT GGCGAATGCT   120
 8 GTGGTAGTTA GGCCTTTTCT TTCTCACCAG CAGATTGAGA TCCTCATTAA CCTAATGCAA   180
 0 CGTGGTAGC TTGTTTTCGG GGCCGAGGTT TTCTGGAATC ATCCCATCCA GCGTGTCATC   240
 2 GATAACAGC TGGAGCTTTA CTGCCGCGCC CGCTCCGGCC GCTGTCTTGA AATTGGCGCC   300
 4 CATAACAGC TGGAGCTTTA CTGCCGCGCC CGCTCCGGCC GCTGTCTTGA AATTGGCGCC   360
 4 CATCCCGGCT CAATAAATGA TAATCCTAAT GTGGTCGACC GCTGCTTCCT CCGCCCTGTT   420
 6 GGGCGTGATG TTCAGCGCTG GTATACTGCT CCGACCGCG GGCCGGCTGC TAATTGCCGG   480
 8 GTTCTGCGC TGGGCGGCCT TCCGGCTGCT GAGTAATTGT ACTTCCTTGA CGGGTTTTCT   540
 0 GGGTGTAACT TTCCCGGTCG GAGTGGCATG GATTTGTATT GCGTTCATGA TATGTGAGTA   540
```

file://C:\Crf3\Outhold\Vsrl851410.htm

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/851,410

DATE: [illegible]
TIME: [illegible]

Input Set : **N:\Crf3\RULE60\09851410.txt**
Output Set: **N:\CRF3\06152001\I851410.raw**

```
L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:50 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:251 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3
L:273 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4
L:294 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5
L:367 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:367 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=6
L:361 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:1142 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1142 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=10
L:1135 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10
L:1405 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1405 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=11
L:1346 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11
L:1477 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1477 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=12
L:1470 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12
L:1651 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1651 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=17
L:1680 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1680 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=18
L:1708 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1708 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=19
L:1736 M:220 C: Keyword misspelled or invalid format, [(C) STRANDEDNESS:]
L:1736 M:246 W: Invalid value of Alpha Sequence Header Field, [STRANDEDNESS:], SeqNo=20
```